UNCERTAINTY IN EAGLE FORD SHALE WELLS PRODUCTION QUANTIFIED USING PROBABILISTIC ANALYSIS WITH A NOVEL DECLINE-CURVE MODEL

A Thesis

by

ISAAC DAVID ZHUKOVSKY

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Michael J. King
Co-Chair of Committee,	Akhil Datta-Gupta
Committee Member,	Mark E. Everett
Head of Department,	A. Daniel Hill

May 2017

Major Subject: Petroleum Engineering

ABSTRACT

This thesis proposes a Bayesian decline curve methodology, using Markov Chain Monte Carlo (MCMC) simulation and a novel empirical decline curve equation to better quantify uncertainty in estimated ultimate recovery (EUR) for oil shales. The methodology is calibrated using hindcasting of production data from an area of the Eagle Ford oil window. Hindcasting on an areal basis with 254 wells has demonstrated good results, with a coverage rate of true reserves of 78% for an 80% confidence interval (P90-P10), or 199 of the 254 wells tested. The novelty of the new model is in the implementation of the empirical decline curve equation for shale wells in Bayesian decline curve analysis with fast per well solution time on typical engineering computers

This method offers many benefits. Principally, the method quantitatively assesses uncertainty and avoids subjective estimates of uncertainty. The results the method generates are accurate for shale wells because the decline curve equation was empirically designed for such wells. Furthermore, it generates replicable results for given wells regardless of the forecasting engineer and offers a fast calculation time of 5-10 seconds per well in the data set.

The novel decline curve equation used in this methodology accommodates both early steep rate decline and later shallower decline with a smooth transition. Coupled with a Bayesian decline curve analysis process, the decline behavior of shale wells is assessed probabilistically with accuracy. The decline curve parameters are random variables with defined prior distributions. A MCMC simulation is performed to obtain the posterior distribution of well EURs. Wells in the sample set had >36 months of production with the first 12 months used as simulation input. The method is calibrated (hindcasted) on an areal basis by measuring the coverage rate of true reserves. The method, when applied to the

Eagle Ford oil shale production data, demonstrates good convergence to stationary posterior distributions of the parameters. This is important because field development of the Eagle Ford and other shale plays is improved when uncertainty is accurately quantified. By quantifying uncertainty and moving away from deterministic decline curve analysis using equations designed for conventional reservoirs, a better understanding of shale well EUR and behavior throughout a play is obtained.

DEDICATION

To my family, Donna, William, Sarah, Lucy, and Wash, and my friends for all their support and encouragement.

ACKNOWLEDGMENTS

I would like to thank Dr. Michael King, my committee chair, for all his support and guidance during my research. My sincere thanks to Dr. Akhil Datta-Gupta and Dr. Mark Everett for their support and guidance during my research and for generoulsy agreeing to serve on my committee.

I also want to thank all of my colleagues in the Model Calibration and Efficient Reservoir Imaging (MCERI) group for acting as sounding boards, for giving advice, and for being great people to work with during my years at Texas A&M.

Special thanks go out to Phaedra Hopcus. She has always been willing to help me navigate any paperwork or administrative obstacle I encountered, and I appreciate her continued dedication to making sure MCERI is organized and running smoothly.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a thesis committee consisting of Professor Michael King (advisor) and Professor Akhil Datta-Gupta (co-advisor) of the Department of Petroleum Engineering and Professor Mark Everett of the Department of Geology and Geophysics.

Tuning of the proposal distributions and portions of the data filtering process were performed collaboratively with Ruben Mendoza of the Department of Petroleum Engineering and were published in 2016 in an article listed in the references section. All other work conducted for the thesis was completed by the student independently.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University and a thesis research assistantship from the CMG Foundation.

NOMENCLATURE

CDF	Cumulative Distribution Function
EUR	Estimated Ultimate Recovery
k	Permeability
PDF	Probability Density Function

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

I. INTRODUCTION*

1.1 Objectives

The objectives of this work are:

- To *propose* a new method for quantifying uncertainty in Eagle Ford Shale wells using a Bayesian framework
- To *demonstrate* the new method on multiple wells within the Eagle Ford Shale
- To *validate* the new method by performing error analysis

1.2 Statement of the Problem

One of the most commonly applied models in the field of decline curve analysis is Arps's (1945) Rate-Time equation. The equation is an empirical construction developed for wells producing under constant bottom hole pressure in conventional reservoirs. These wells typically have permeabilities of milidarcies or greater, as contrasted with the nano-darcy scale permeability encountered in shale reservoirs. Furthermore, Arps (1945) equation is for wells that have not undergone hydraulic fracturing, as opposed to shale wells which are almost always the subject of multi-stage hydraulic fracturing treatments. While Arps' equation can be applied to hydraulically fractured wells, it is still a departure from the conditions of the original empirical construction, which assumes boundary dominated flow. Rushing et al. (2007) conducted a study which found that the largest source of error was the incorrect application of the Arps equation during the transient or transitional flow

periods. These flow regimes tend to dominate the early and middle life of shale wells, complicating accurate forecasting using Arps' equation. Many new decline curves have been proposed for shale wells in recent years (Duong 2011; Valko and Lee 2010; Valko 2009; Anderson et al. 2010; Ilk et al. 2008). The principles for some of these decline curves are briefly summarized below:

- Non-Hyperbolic power law loss ratio formulation (Ilk et al. 2008)
  - Requires determination of decline rate at infinite time, D_∞, which can be difficult to acquire using blind statistical methods (Johnson et al. 2009)
- Composite method that estimates contributions from the Stimulated Reservoir Volume and unstimulated matrix (Anderson et al. 2010)
  - If boundary dominated flow is not observed at time of analysis, requires use of microseismic data or well spacing to infer size of stimulated reservoir volume
- Stretched Exponential Decline as an approximation of the behavior of a great number of contributing volumes in exponential decay (Valko 2009; Valko and Lee 2010)
  - The empirical model parameter n must be determined iteratively
  - The Stretched Exponential Decline curve assumes a terminal production rate of zero
- Fracture Dominated Flow for life of well (Duong 2011)
  - Does not deal well with flow regime changes (Freeborn and Russell 2012)
  - Requires terminal production rate to be zero

The other major source of difficulty in applying decline curves to shale wells stems from the almost exclusive use of deterministic analysis. According to Lee and Sidle (2010)

uncertainty analysis of production forecasts could provide valuable insight into the upside and downside potential of resources plays and could allow for more accurate categorization of reserves. Some early attempts at probabilistic reserves and uncertainty quantification include the bootstrap method, developed by Jochen and Spivey (1996), which involves generating synthetic data sets using Monte Carlo simulation of the production data and applying non-linear regression to determine many sets of decline curve parameters. While a step in the right direction, the bootstrap method requires a sufficient number of production data points to generate a meaningful quantity of synthetic data sets and may require more iterations than a directed form of Monte Carlo search. Gong et al. (2011) presented a Bayesian framework for probabilistic decline curve analysis using Arps's (1945) equation in the Barnett shale for 167 wells with more than 7 years of production history. This method demonstrates good coverage of the true reserves while offering a calculation time 10 to 13 times faster than the bootstrap method (Gong et al. 2011).

We present a Bayesian method that combines Gong et al.'s (2011) Bayesian framework with Zhang et al.'s (2015) decline curve empirically derived for shale wells after Zhukovsky et al. (2016). Zhang et al. (2015) derived their decline curve by modifying the exponential form of Arps's (1945) as in **Eq 1.1**. Zhang et al. (2015) focused on the concept of a growing drainage volume in time where the a term in **Eq 1.1** is replaced with a more complex expression as in as in **Eq 1.2**.

$$q(t) = q_i \exp^{-at} \tag{1.1}$$

$$a = \beta_l + \beta_e \exp^{-t^n} \tag{1.2}$$

Eq 1.2 consists of two decline rates: an early time component, β_e, that decays exponentially in time and a late time component, β_l, that becomes more dominant as β_e decays. β_e signifies the early life, transient period of shale wells, while β_l represents the decline

rate of the subsequent flow regime, likely that of full fracture interference. The n is an empirical exponent that governs when the smooth transition from the early decline trend to the late decline trend occurs. From their examination of more than 100 Eagle Ford shale wells, Zhang et al. (2015) determine that the n factor ranges from 0 to 0.7. More information on the parameters is displayed in **Table 1.1**.

Table 1.1: Decline curve parameters and descriptions for Zhang et al.'s (2015) equation

Parameter	Description
β_e	Decline constant to account for the early time, fully transient period
β_l	Decline constant to account for late life period
n	Empirical exponent with a recommended range of 0 to 0.7
t	The time in months
q_i	The initial production rate

We start with the Bayesian Markov Chain Monte Carlo methodology developed by Gong et al. (2011) and develop it as in Zhukovsky et al. (2016). By using Bayes' rule, **Eq 1.3**, we develop a directed MCMC search workflow that generates an entire chain of decline curves for each Eagle Ford shale well, as well as distributions of the decline curve parameters, EUR, production rates at each month, and other parameters of interest. In this framework, we choose deliberately broad distributions of the decline curve parameters in order to explore the space of reasonably possible decline curves (Zhukovsky et al. (2016)) and let the production data serve as the update that gives the posterior distributions of the decline curve parameters their shape. **Eq. 1.3** governs the relationship between the prior distributions and the updating information.

$$P(A|B) = \frac{P(B|A)P(B)}{P(B)} \tag{1.3}$$

In practice, it is often difficult to obtain all of the distributions in Bayes' rule directly, which is why we use the iterative Markov Chain Monte Carlo method. When examining proposed samples to the Markov Chain, we calculate the acceptance ratio, **Eq. 1.4**, and compare it against a random variate between 0 and 1 in order to decide whether to accept or reject the sample.

$$\alpha = \min\left(1, \frac{\exp(-\frac{\sigma^2_{proposal}}{\sigma^2})}{\exp(-\frac{\sigma^2_{s-1}}{\sigma^2})} * \Pi_{v=k_{r,o},k_{r,w}} \frac{\Phi\left(\frac{v_{upper}-v_{s-1}}{\sigma_v}\right) - \Phi\left(\frac{v_{lower}-v_{s-1}}{\sigma_v}\right)}{\Phi\left(\frac{v_{upper}-v_{proposal}}{\sigma_v}\right) - \Phi\left(\frac{v_{lower}-v_{proposal}}{\sigma_v}\right)}\right) \tag{1.4}$$

An acceptance ratio greater than 1, indicating a proposed set of decline curve more likely than the previous sample from the $s-1$ step in the chain, is set to 1 and automatically accepted as the next step in the chain. Samples that are less likely, with larger misfit, still have a chance to be accepted as part of the chain. In this manner, Markov Chain Monte Carlo methods, and the Metropolis-Hastings algorithm which we employ here, explore the high likelihood regions of the posterior distribution while still maintaining the ability to jump out of local minima. The Markov Chain Monte Carlo workflow follows **Fig. 1.1** (Zhukovsky et al. 2016).

1.3 Results and Validation

In order to validate the method, we apply the workflow and use it to forecast a single Eagle Ford shale well, the Warren B well. We acquire the production data from DrillingInfo and load it into excel, identifying the oil, gas, and water rates as available. We use excel's solver function to perform a least squares fit according to the constraints in **Table 1.2**. These constraints are chosen in order to allow the solver to explore the parameter space of plausible (and physically consistent) combinations. Once the least squares fit for the parameters is obtained, we implement Markov Chain Monte Carlo sampling by

reimplementing the Bayesian algorithm presented in Gong et al. (2011) in Matlab.

Table 1.2: Solver constraints for decline curve parameters

Parameter	Constraint
β_e	0.01 <= β_e <= 1.5
β_l	0.01 <= β_l <= 0.15
n	0 <= n <= 2.5
q_i	0 <q_i <3,300,000 STB/mo

The prior distributions are chosen in order to allow method to explore many plausible decline curve parameter combinations. We attempt to avoid overly constraining the Markov Chain Monte Carlo search. The method is allowed to proceed for 100,000 iterations. Then, output figures are generated for the EUR of the well (**Fig. 1.2**) and distributions are calculated for the forecasted production at each month. The distributions of forecasted production are used to generate the probabilistic decline curves for the well as shown in **Fig. 1.3**.

We have demonstrated that the method provides reasonable estimates of production for the Warren B well. Uncertainty is quantified via the method and represented by the spread between the P90 and P10 curves. In this particular case, it turns out that the P90 case was more accurate by observing the production data that is held back from the method and not used as input.

Figure 1.1: The Markov Chain Monte Carlo Workflow: Wells Are Processed Sequentially and the Markov Chain Monte Carlo Simulation is Performed for the First 12 Months of Oil Production, the Rest of the Production Data is Used for Hindcasting.

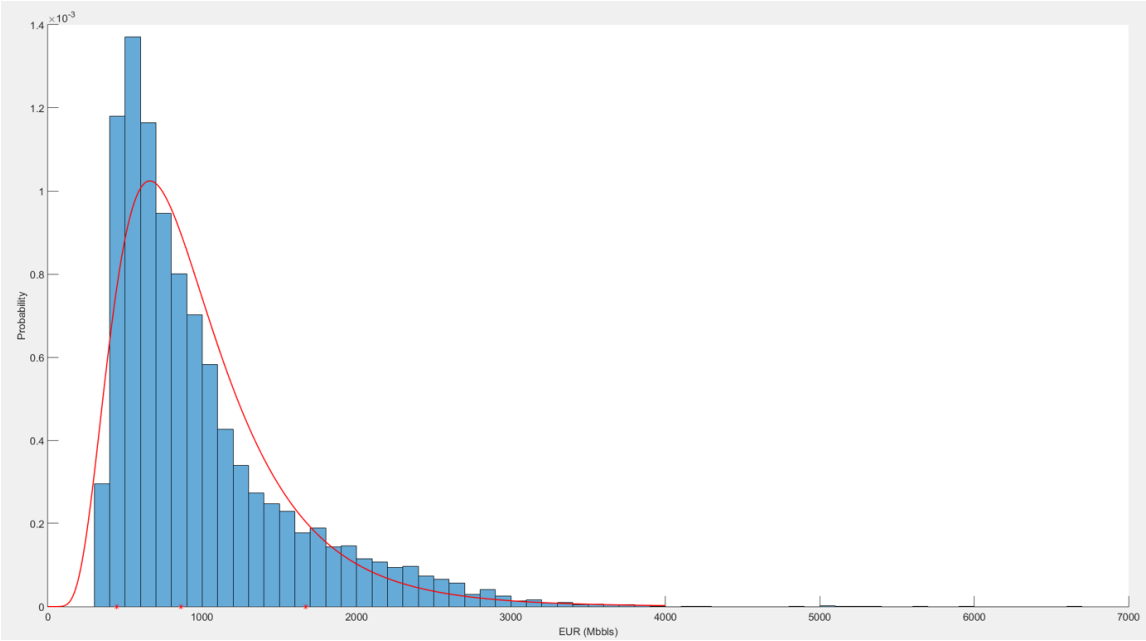

Figure 1.2: Distribution of EURs for the Warren B Well in the Eagle Ford. The P90, P50, and P10 Are Denoted in Sequence by the Red Stars.

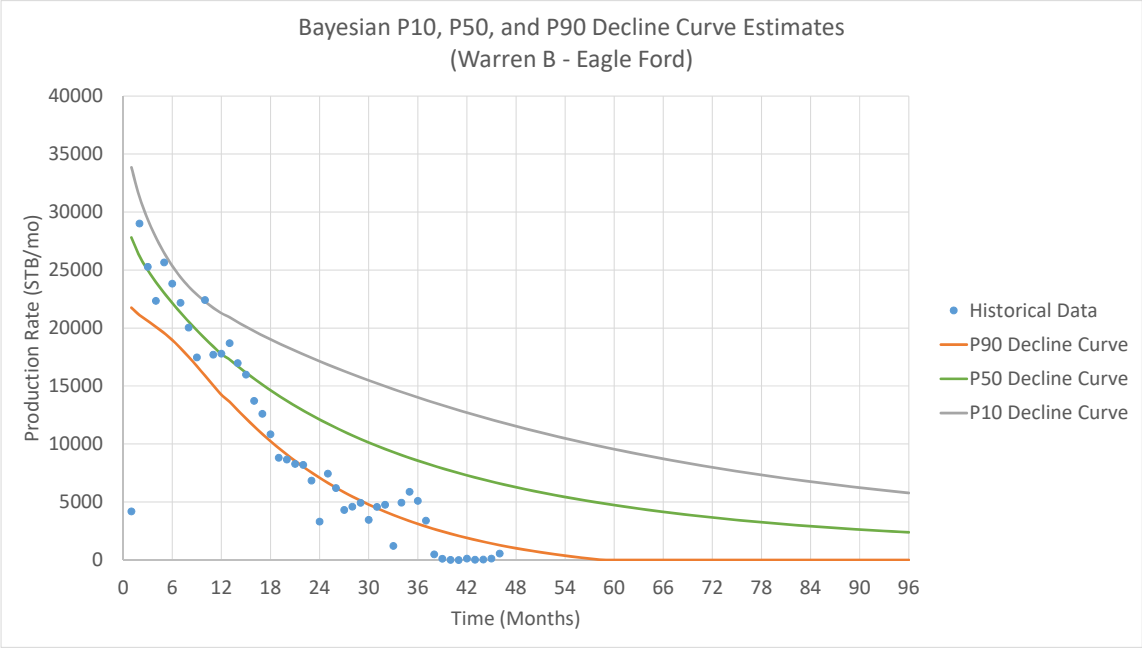

Figure 1.3: Bayesian Decline Curves for the Warren B Well. The First 12 Months of Data is Used as Input and We Can See That the P90 and P10 Curves Bracket the Production During the Input Phase.

II. LITERATURE REVIEW - DECLINE CURVE ANALYSIS IN SHALE WELLS

In this chapter, we review the history of decline curve analysis in shale wells. We divide the literature review into two sections: a review of the decline curve models, which have a long and well documented history of application to many types of reservoirs, and probabilistic methods as applied to decline curve analysis, a rather more recent development. The shale boom is a relatively recent phenomenon, but that has not stopped petroleum engineers from applying tried and true decline curve models and developing new ones in order to evaluate production from shale wells. We discuss both classic models and novel models as applied to shale reservoirs in our literature review. Furthermore, the higher uncertainty inherent in shale reservoirs and decline curve evaluation has spurred renewed focus on probabilistic methods that quantify uncertainty, unlike deterministic methods that only provide qualitative measures of uncertainty. Probabilistic methods have been implemented using traditional decline curve models and novel models and we discuss both implementations.

2.1 Review of Decline Curve Models

We begin our review of decline curve analysis by examining the three models presented in Arps 1945: exponential decline (**Eqs. 2.1, 2.2**), hyperbolic decline (**Eqs. 2.3, 2.4**), and harmonic decline (**Eqs. 2.5, 2.6**). The variables present in the three models are explained in **Table 2.1**. Exponential decline was empirically derived under the following assumptions:

- Idealized reservoir
  - No aquifer drive present
  - Pressure of reservoir is proportional to amount of remaining oil

- Productivity indexes of wells are constant through well life so production rates are always proportional to reservoir pressure

$$q = q_i \exp^{-d_i t} \tag{2.1}$$

$$N_p = \frac{q_i - q}{d_i} \tag{2.2}$$

While a reasonable and straightforward relationship, exponential decline is not the most commonly observed behavior in conventional oil reservoirs. More typically, pressure tends to decline at a lower rate as more oil is produced, leading to the hyperbolic decline model (**Eqs. 2.3, 2.4**). This model is one of the most frequently applied models in both conventional reservoirs, hydraulically fractured horizontal wells, and unconventional shale reservoirs (which are frequently developed using hydraulically fractured horizontal wells). Originally, exponential decline was used as a (low) approximation to the more common hyperbolic decline since its simpler mathematical form made it much easier to work with. More recently, certain shale plays have exhibited sharp initial declines that may appear exponential on paper, but rarely do they exhibit exponential decline for the whole well life.

$$q = \frac{q_i}{(1 + b d_i t)^{\frac{1}{b}}} \tag{2.3}$$

$$N_p = \left[\frac{a_i^b}{(b-1)d_i}\right] * \left[q^{(1-b)} - q_i^{(1-b)}\right] \tag{2.4}$$

Hyperbolic decline is considered the most general case of the three models presented (Arps 1945). Unlike with exponential decline, the effective decline rate changes over time in hyperbolic decline. This results in a gradual decrease of effective decline rate

from the initial value which results in the characteristic leveling off of the production rate at later time. This behavior is commonly referred to as the hyperbolic tail or "tailing off." In the original hyperbolic decline construction, the b factor is constrained to between 0 and 1. If the b factor is 0, the equation simplifies to exponential decline. If the b factor is 1, the equation simplifies to harmonic decline (**Eqs. 2.3, 2.4**). For the original model, b factors above 1 are not permitted, but in practice b factors greater than 1 can be seen in hydraulically fractured wells. Care must be taken when using the hyperbolic decline equation with a b factor greater than 1; unless a minimum terminal decline rate or abandonment rate is specified, it can be demonstrated that the cumulative production equation converges to infinity.

$$q = \frac{q_i}{(1 + d_i t)} \tag{2.5}$$

$$N_p = \left[\frac{q_i}{d_i}\right] * \ln\left[\frac{q_i}{q}\right] \tag{2.6}$$

The last form of Arps's (1945) decline curve equation we discuss is the harmonic decline model. It is a special case of hyperbolic decline, as previously mentioned, and is also the model with the highest b factor that still produces a finite EUR. In conventional reservoirs, the harmonic estimate is considered the most optimistic of the three models. In unconventional reservoirs, harmonic decline is not considered particularly noteworthy since the wells are usually hydraulically fractured with b factors greater than 1.

While Arps's (1945) model is fairly versatile , it cannot be applied during transient flow regimes while still honoring the original assumptions of the model. Incorrectly applying these equations during transient flow can lead to large errors in predicted EUR (Rushing et al. 2007). The need for better empirical and mechanistic decline models for shale wells led to new models, such as the work of Ilk et al. (2008). The authors modify the original

Table 2.1: Decline curve parameters and descriptions for Arps's (1945) decline models

Parameter	Description
q	Current production rate
q_i	Initial production rate
d_i	Initial nominal decline rate at $t = 0$
t	Cumulative time since start of production
N_p	Cumulative production
b	Hyperbolic decline constant ($0 < b < 1$)

Table 2.2: Decline curve parameters and descriptions for Ilk et al.'s (2008) "Power Law Loss Ratio" decline model

Parameter	Description
q	Current production rate
$\hat{q}_i$	Rate "intercept" [*i.e.* $q(t = 0)$]
D_∞	Decline constant at "infinite time" [*i.e.* $D(t = \infty)$],
$\hat{D}_i$	Normalized initial decline constant [*i.e.* $\hat{D}_i = D(t = 1timeunit)/n$],
t	Time
n	Time exponent

equation to account for the relationships that the b factor and loss ratio, D, have with time. The authors determine that the two parameters have a power law relationship with time, leading to their "Power Law Loss Ratio" Rate decline model (**Eq. 2.7**).

$$q = q_i \exp[-\hat{D}_\infty t - \hat{D}_i t^n] \tag{2.7}$$

While the relationship may not be obvious when examining **Eq. 2.7**, $\hat{D}_i$ and $\hat{D}_\infty$ can be combined into a composite decline factor as in **Eq. 2.8**.

$$D = D_\infty + D(t = 1)t^{-(1-n)} \tag{2.8}$$

This composite decline factor embodies an idea that also appears in later work: to

account for early transient flow and later stabilized flow, composite factors should be used. Also, at late time, D_∞ dominates and the power law equations essentially behave like the exponential decline equations. While a step in the right direction for decline analysis in shale wells, Ilk et al.'s (2008) "Power Law Loss Ratio" Decline Model has some limitations and drawbacks:

- Requires use of diagnostic plot to determine various D parameters
- D_∞ must be calibrated correctly, and this may require more data than is available at earlier times
- Blind statistical methods cannot be used to determine the power law parameters (Johnson et al. 2009)

Different work presents the model of Stretched Exponential Decline, formulated on a completely empirical basis (Valko 2009; Valko and Lee 2010). The Stretched Exponential model requires the evaluation of both 1 and 2 parameter Gamma functions or, alternatively, the incomplete Gamma function, making the model somewhat unique in that regard. The authors suggest the following procedure for performing single well analysis using the Stretched Exponential Decline Model:

1. Prepare a data series of Q and q_d (dimensionless production rate)
2. Assume a value for the n parameter and calculate rp (recovery potential - **Eq. 2.9**)
3. Plot of rp vs Q (cumulative production - **Eq. 2.10**) should appear as a straight line
4. If the intercept of the actual data is not 1, adjust the n value and repeat from step 2

$$rp = 1 - \frac{Q}{\text{EUR}} = \frac{1}{\Gamma\left[\frac{1}{n}\right]}\Gamma\left[\frac{1}{n}, -\ln\frac{q}{q_0}\right] \tag{2.9}$$

$$Q = \frac{q_0 \tau}{n} \left\{ \Gamma \left[\frac{1}{n} \right] - \Gamma \left[\frac{1}{n}, \frac{t}{\tau}^n \right] \right\} \tag{2.10}$$

$$q = q_0 \exp \left[-\frac{t}{\tau}^n \right] \tag{2.11}$$

The Stretched Exponential Decline Model analysis process can result in accurate estimations of EUR. Moreover, it is formulated with a good empirical understanding of how shale gas data behaves. Unfortunately, it assumes that the terminal production rate is 0 which can lead to inflated EURs when economic constraints mandate a significantly non-zero abandonment rate, and the determination of the n parameter is an iterative process.

Additional approaches have used a more theoretical, rather than empirical, basis such as the work of Johnson et al. (2009). The authors observe that linear flow is typically the dominant flow regime as the stimulated reservoir volume of a shale well is depleted and that many previous approaches, such as Arps (1945), are biased towards boundary dominated flow. Consequently, the authors propose a production analysis workflow that produces 3 (deterministic) estimates of EUR. The approach uses pressure transient analysis to determine the properties and dimensions of the stimulated reservoir volume. Johnson et al. (2009) use the following work flow:

1. Specify appropriate input parameters and constraints
    - Ex: horizontal well length, fracture height, etc.
2. Using pressure transient analysis and a log-log plot, identify flow regimes
    - Ex: Linear flow, boundary dominated flow, etc.
3. If boundary dominated flow is observed, determine hydrocarbon pore volume and calculate stimulated reservoir width

4. Using the square-root time plot, determine the slope and intercept and use the parameters to calculate the apparent skin and the linear flow parameter

5. Calculate the volume of the stimulated reservoir volume using the parameters in hand

    - If boundary dominated flow was not observed, use an external observation to determine the stimulated reservoir width, such as microseismic or well spacing

6. Now, use a preferred analysis model to calculate an EUR for the stimulated reservoir volume only, the stimulated reservoir volume with minimal matrix contribution ($k \sim 1\text{e}{-6}$), and the stimulated reservoir volume with substantial matrix contribution ($k \sim 1\text{e}{-4}$)

It is important to note that while this method produces three EURs, it is purely deterministic and not probabilistic. Furthermore, at early times when boundary dominated flow has not been observed, it requires the use of microseismic data or well spacing to determine the size of the stimulated reservoir volume which is less than ideal. Finally, while it can be used with both simulations or analytical models and is somewhat model agnostic, use of complex models in the final step of the work flow can add considerably to the time required to perform the analysis.

A less complicated implementation for the evaluation of early time linear flow is proposed by Duong (2011). The author acknowledges the presence of prolonged linear flow in shale wells and provides analysis of several long production shale gas and shale oil wells to demonstrate the domination of the linear flow regime throughout well life. He proposes that for wells in linear flow, the dimensionless decline rate follows **Eq. 2.12.**

$$\frac{q}{N_p} = at^{-m} \tag{2.12}$$

This equation can be transformed into a more useful form for individual well analysis (**Eqs. 2.13, 2.14**; Parameter descriptions in **Table 2.3**).

$$q = q_1 t^{-m} \exp^{\frac{a}{1-m}(t^{1-m}-1)} \tag{2.13}$$

$$q = q_1 t^{-m} \exp^{\frac{a}{1-m}(t^{1-m}-1)} + q_\infty \tag{2.14}$$

$$t(a,m) = t^{-m} \exp^{\frac{a}{1-m}(t^{1-m}-1)} \tag{2.15}$$

Table 2.3: Decline curve parameters and descriptions for Duong's (2011) decline model

Parameter	Description
q	Current production rate
$\hat{q}_1$	Flow rate at Day 1
a	Intercept of line fit of production on log-log plot
m	Slope of line fit of production on log-log plot
t	Time in days
q_∞	Flow rate at infinite time (Can be positive, negative, or 0)
q_{eco}	Economic cutoff for flow rate
t_{eco}	time where $q = q_{\text{eco}}$

Duong (2011) proposes the following work flow:

1. Clean and format the production data
2. Plot the data on a log-log plot and fit a straight line to determine a and m
    - Y axis: $\frac{q}{N_p}$, Current production/cumulative production
    - X axis: Time, days

3. Create q vs $t(a,m)$ plot (see **Eq. 2.15** for $t(a,m)$) and fit a straight line to determine q_1

4. Apply forecasted rates to rate-time or rate-cum. plot for validation

5. Calculate EUR if desired (See **Eq. 2.16**)

$$\mathrm{EUR} = \frac{q_{\mathrm{eco}}}{a} t_{\mathrm{eco}}^{m} \tag{2.16}$$

Duong's (2011) model has proven quite popular since its introduction; it has even been implemented into several industry standard well economics programs. There are a few drawbacks, however:

- Determining a and m requires the use of a log-log diagnostic plot
- Determining q_1 requires a specialized diagnostic plot
- The method assumes the terminal rate is 0 when determining key decline curve parameters
- The method has difficulty with changing flow regime (Freeborn and Russell 2012)

Finally, we examine the work of Zhang et al. (2015). Based on the idea of "growing drainage volume" the Zhang et al. (2015) model accommodates both transient (usually linear) flow and the boundary dominated flow that follows, with a smooth transition between the curves for the two flow regimes. In fact, of the curves we have reviewed, this model is the only one that easily accommodates two different flow regimes with a tunable, smooth switching point. The model is based on a modification of the equation Fetkovich (1980) presented for Arps' style decline under the assumptions of constant bottom-hole pressure and boundary dominated flow (**Eq. 2.17**).

$$q(t) = q_i \exp^{-at} \tag{2.17}$$

Zhang et al. (2015) modify the a factor into a composite of two decline components with a decaying exponential factor (**Eq. 2.18**).

$$a = \beta_l + \beta_e \exp^{-t^n} \tag{2.18}$$

Early on, flow from the fracture network is the dominant contribution to shale well production and is represented by the β_e term. As the fracture network is drained, the $\exp^{-t^n}$ term, of **Eq. 2.18** decreases the effect of the β_e decline constant and increases the relative importance of the β_l or late time decline factor. The n term governs when the method switches from the transient decline behavior to the late-time decline behavior. Once the depletion of the fractures results in a drainage volume for each fracture that can "see" the boundary of the drainage volume of neighboring fractures, full fracture interference has been achieved and the late-time decline factor, β_l dominates. Zhang et al.'s (2015) model compares favorably to other models in the literature. An example comparison is given in **Fig. 2.1**. An added advantage of the method is that it does not require specialized plots to use, nor does it require data external to the production history. Although it is a fully empirical model, it attempts to capture the underlying behavior of the flow regimes in shale reservoirs through its construction.

There are many decline curve models available for shale gas wells, some more suited the task than others. From the venerable Arps' models to the far more recent Duong method, the available models vary in complexity and accuracy. We feel the best balance between complexity and accuracy is Zhang's model, and we incorporate it into our probabilistic decline method as we describe in Chapter III.

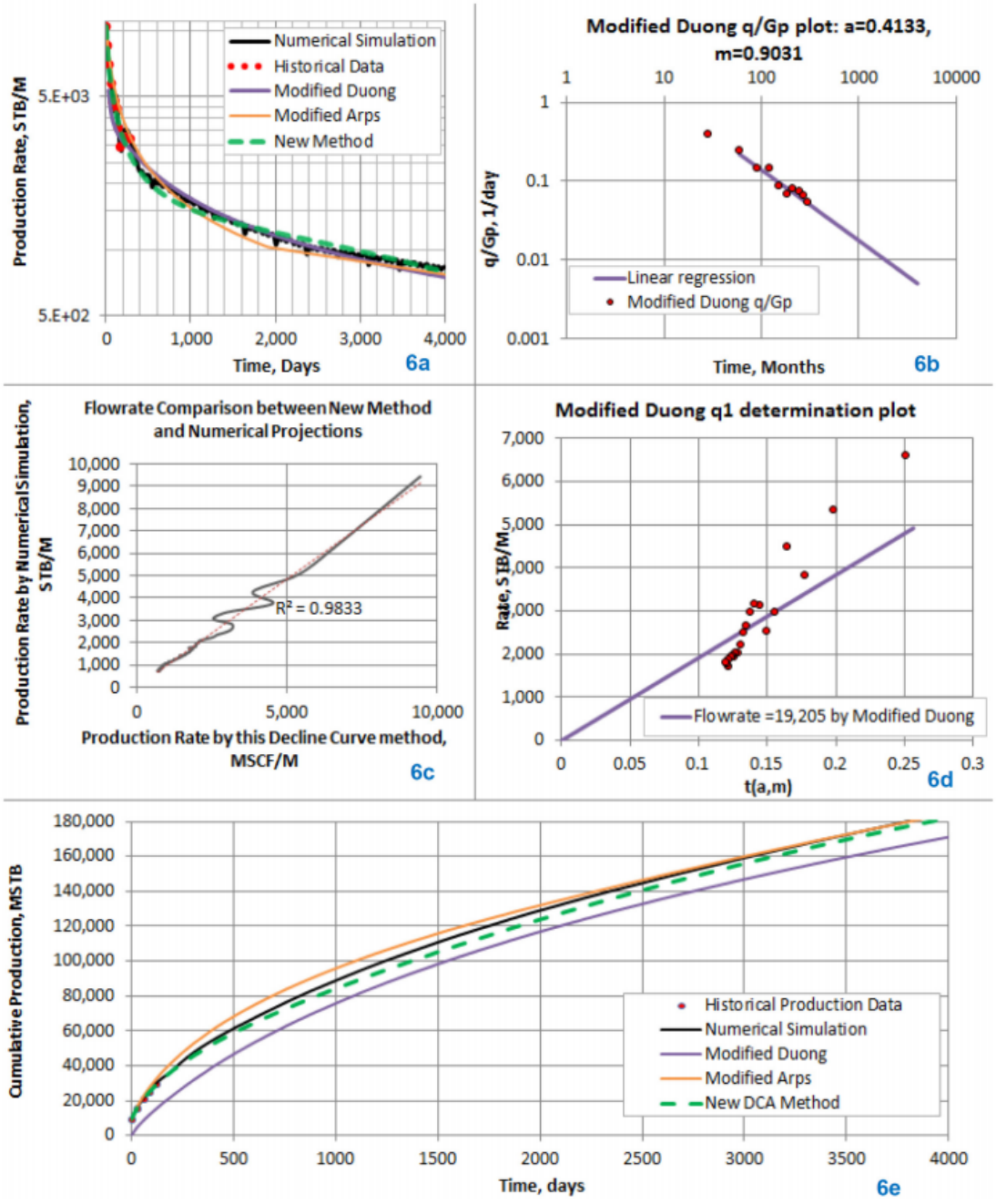

Figure 2.1: Comparison of Various Decline Models for an Eagle Ford Shale Oil Well Producing in Karnes County, Texas. Considering How Closely Each Curve Follows the Output of Reservoir Simulation, Zhang's Equation Performs Quite Well (Zhang et al. 2015)

2.2 Review of Probabilistic Methods

We divide our review of probabilistic methods roughly in half. In the first half of our review, we will discuss the work of Fulford et al. (2016, 2015), who used a Bayesian process involving a Markov Chain Monte Carlo simulation of the decline curve parameters of the Transient Hyperbolic Model (Fulford and Blasingame 2013). In the second half, we discuss the Bayesian methodology of Gong et al. (2011). The authors use a Bayesian Markov Chain Monte Carlo simulation that utilizes Arps's (1945) hyperbolic model, but the framework itself is model agnostic.

Fulford et al. (2016, 2015) develop their method to perform Markov Chain Monte Carlo simulation on the decline curve parameters of the Transient Hyperbolic Model developed by Fulford and Blasingame (2013). The Transient Hyperbolic Model is a modification of Arps's (1945) hyperbolic decline model that involves converting the b factor from a constant into a function that varies with time. The equations for the Transient Hyperbolic Model are presented in **Eqs. 2.19, 2.20, 2.21,** and **Table 3.1**.

$$q = q_i \frac{1}{(1 + bD_i t)^{\frac{1}{b}}} \tag{2.19}$$

$$b(t) = b_{\max} - (b_{\max} - b_{\min}) \exp[-\exp[-c(t - t_{elf}) + \exp[\gamma]]] \tag{2.20}$$

$$c = \frac{\exp[\gamma]}{1.5 t_{elf}} \tag{2.21}$$

Fulford et al. (2016, 2015) based their Markov Chain Monte Carlo simulation around four of the Transient Hyperbolic Model parameters: q_i, D_i, b_f, and t_{elf}. b_i is always assumed to be 2 for transient flow. Non-informative constrained prior distributions are used for q_i and D_i. b_f and t_{elf} have uniform or normal distributions that are defined by the prior

Table 2.4: Decline curve parameters and descriptions for Fulford and Blasingame's (2013) Transient Hyperbolic Model

Parameter	Description
$b_{\max}$	Maximum hyperbolic coefficient [constant b parameter at early-time linear flow regime]
$b_{\min}$	Minimum hyperbolic coefficient [constant b parameter at boundary-dominated flow regime]
t_{elf}	Time to end of linear flow [beginning of transition from $b_{\max}$ to $b_{\min}$]
γ	Euler-Mascheroni constant [0.57721566...]
c	Scaling factor [controls the rate of transition from $b_{\max}$ to $b_{\min}$]
q_i	Initial flow rate
t	Time
D_i	Initial decline rate

belief of the evaluator. The normal distributions, when used in place of uniform distributions, are based around a maximum likelihood estimate of the evaluator. The proposal distributions for each of the parameters is generally on constructed on a log basis, with the exception of the proposal distribution of the b_f factor. For an example of one of the proposal distributions, see **Eq. 2.22**.

$$\ln(q_{i_{\text{proposal}}}) \approx \mathcal{N}[\ln(q_{i_{n-1}}), 0.2] \tag{2.22}$$

Proposal misfit is calculated using both logarithm residuals and the absolute magnitude of the logarithm residuals. Fulford et al. (2016, 2015) suggest the following workflow for the Markov Chain Monte Carlo simulation for a single well:

1. Estimate prior distributions for the parameters q_i, D_i, b_f and t_{elf}
2. Set stating conditions for the Markov Chain by setting the Transient Hyperbolic Model parameters equal to the mean of their respective prior distributions

3. Calculate the product of the prior distribution function and the likelihood distribution function for the model starting conditions

4. Generate a model proposal by use of the proposal distribution functions (ex: **Eq. 2.22**)

5. Calculate the product of the prior distribution function and the likelihood distribution function for the model proposal

6. Evaluate the acceptance ratio, α

7. Sample a variate from the uniform distribution, $U(0, 1)$

8. If the variate is less than the acceptance ratio, then accept the model proposal. Otherwise, reject the model proposal

9. Repeat steps 4 through 8 until the iteration limit is reached

The authors find that their method provided good accuracy, evaluated through hindcasting, of 136 wells in the Elm Coulee field in the Bakken shale. They also examine the effect of changing the length of historical data available to the method on the accuracy of the resulting forecast as seen in **Fig. 2.2**.

The authors achieve good results for several well populations. For 136 liquid-rich shale wells in the Elm Coulee field, their method reliably hindcasts 5-year cumulative production with as little as 12 months of production history. Expressing an informed bias about b_f improves the fit even further. Even when the evaluating engineer expresses an incorrect prior belief, Fulford et al. (2016, 2015) claim that the method will generally converge to the correct answer so long as there is an amount of evidence proportional to the evaluator's belief. The method is also quite fast as written in the C++ programming language.

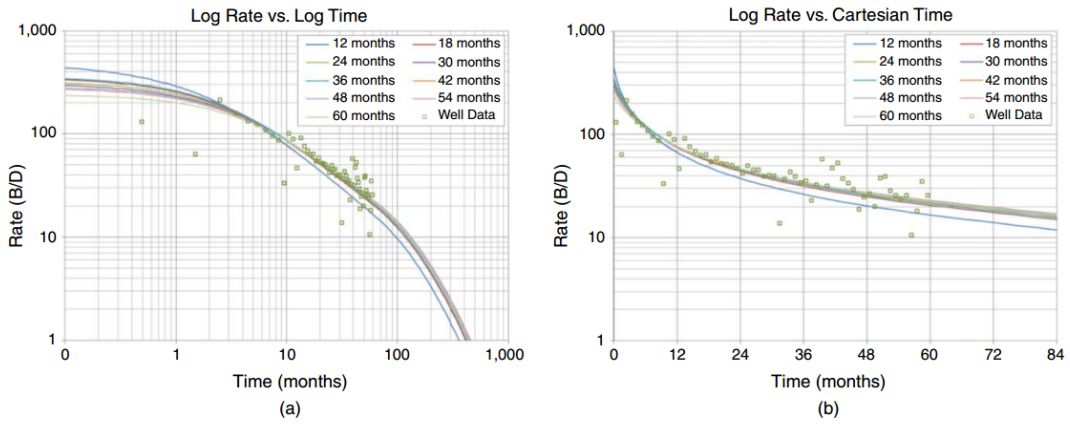

Figure 2.2: Examples of Hindcasts for an Elm Coulee Well with a Known b_f and High Maximum Likelihood Estimate for Varying Length of Historical Data Used to Forecast (Fulford et al. 2016, 2015)

Finally, we turn to the work of Gong et al. (2011). As some of the first pioneers of the use of Markov Chain Monte Carlo methods for decline curve analysis, Gong et al. (2011) are primarily concerned with demonstrating the accuracy and speed of their method as compared to other methods, such as the modified bootstrap method (Cheng et al. 2010), and proving its utility in unconventional plays with limited production data. In their methodology, the authors assume that the decline curve parameters from Arps's (1945) hyperbolic decline curve equation are random variables and use Bayes' Rule (**Eq. 2.23**) and the Metropolis-Hastings algorithm to perform a Markov Chain Monte Carlo simulation. Gong et al. (2011) use non-informative priors in order not to under estimate the uncertainty of the parameters. Their prior-distribution density function can be seen in **Eq. 2.24**.

$$P(A|B) = \frac{P(B|A)P(B)}{P(B)} \tag{2.23}$$

$$\pi[\ln(q_i), \ln(D_i), b] = \frac{1}{18.41 * 6.21 * 2} \tag{2.24}$$

The authors choose priors that do not unnecessarily constrain the Markov Chain Monte Carlo simulation from exploring plausible combinations of parameters, but the priors do constrain the simulation from considering physically inconsistent situations. The authors then filter down to 197 wells in the Barnett shale that exhibit consistent decline trends over prolonged intervals and use these wells to validate their method. The first half of the production data is used to perform the Markov Chain Monte Carlo simulation while the second half of production data is reserved for hindcasting the well. Markov Chain Monte Carlo simulation is performed for all 167 Barnett shale wells and a hindcasting percentage of 85% is achieved. An example of the results of Gong et al.'s (2011) method compared with modified bootstrap method results for a single well can be seen in **Fig. 2.3**. Overall, Gong et al.'s (2011) Bayesian methodology obtains good hindcasting coverage, similar to that of the Modified Bootstrap Method, with a narrower range of P90-P10 EURs. Furthermore, the methodology does so with an order of magnitude reduction in calculation time compared to the Modified Bootstrap Method. Finally, Gong et al. (2011) note that their methodology can be integrated with other analytical decline curve models with only minor changes, paving the way for the methodology introduced in this thesis.

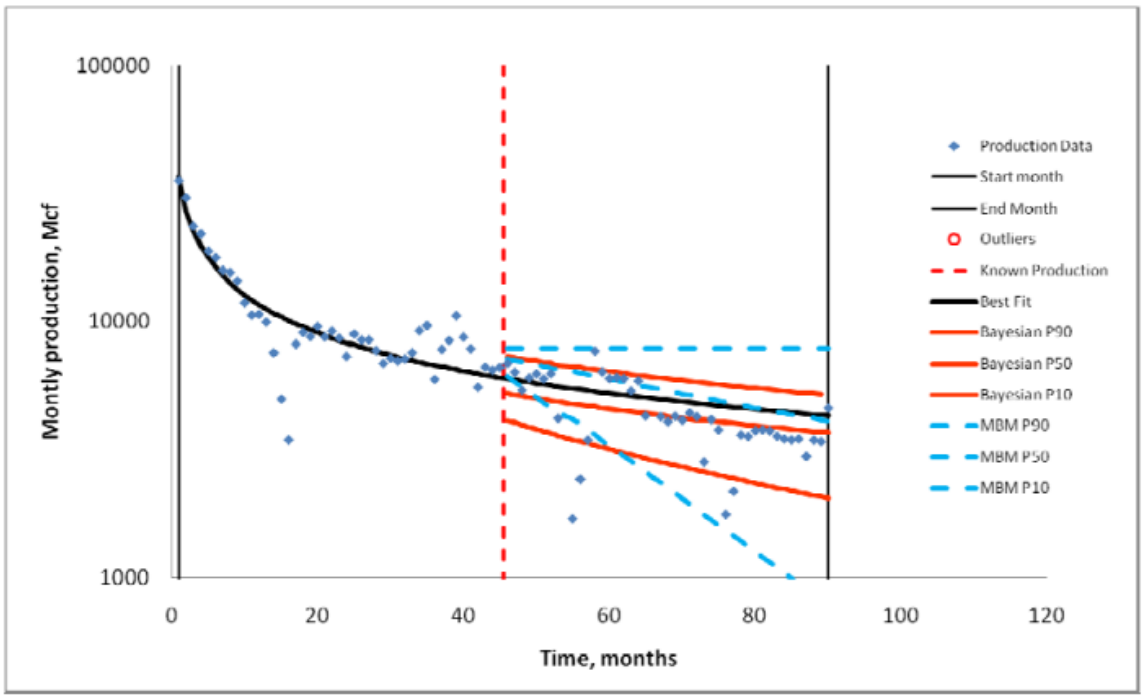

Figure 2.3: Bayesian Forecasts Compared with Modified Bootstrap Method Forecasts for a Single Well in the Barnett shale. The Bayesian Method Produces Estimates that Bracket the Production with a Lower P90-P10 Spread than the Modified Bootstrap Method (Gong et al. 2011)

III. DEVELOPMENT OF THE METHOD*

In this chapter, we discuss the development of the Bayesian decline curve analysis process that implements Zhang et al.'s (2015) equation. Probabilistic methods have been applied in many different aspects of the oil and gas industry for decades. For example, the entire technical specialty of geostatistics is based around using probability theory and its relationships to describe the natural variations of reservoir properties over distance. Panda and Lake (1994) have demonstrated that permeability follows a log normal distribution, enabling engineers to describe permeability using appropriate statitical distributions. The advantage of using a distribution to characterize a parameter of interest lies in its quantification of uncertainty. However, we must intelligently make use of the data we have in order to properly find the distribution of a parameter of interest. One such method for doing so is called Monte Carlo simulation. A Monte Carlo simulation uses random sampling from a distribution, or distributions, in order to predict a large number of outcomes that form a distribution of results (Capen 1976). Given a reasonably small number of contributing distributions, a random walk Monte Carlo simulation stands a good chance of fully characterizing the possible parameter space. However, as the size of the problem expands to include more factors and input distributions, we encounter what is called in statistics the Curse of Dimensionality: as dimensionality increases, the volume of the parameter space increases so quickly that the observed data becomes sparse. In order to overcome this obstacle, we employ a common method of directed search: the Markov Chain Monte Carlo simulation. Markov Chain Monte Carlo algorithms provide methods of obtaining

∗. Part of this chapter is adapted with permission from "Uncertainty Quantification in the EUR of Eagle Ford Shale Wells Using Probabilistic Decline-Curve Analysis with a Novel Model" by Isaac Zhukovsky, Ruben Mendoza, Michael King, and W. John Lee, 2016, Presented at Abu Dhabi International Petroleum Exhibition & Conference, 7-10 November, Abu Dhabi, UAE. Copyright 2016 by Society of Petroleum Engineers.

sequences of random samples from a probability distribution from which direct sampling is difficult. Markov Chain Monte Carlo methods expand upon Bayes Rule (**Eqs. 1.3 and 2.23**) in order to create a stochastic chain process that approximates the posterior probability distribution (Zhukovsky et al. 2016). Furthermore, it can only be considered a Markov Chain process if it has a limited memory of one step: each new sample is dependent on only the step in the chain before it (**Eq. 3.1**).

$$P(X_s|X_1 = x_1, X_2 = x_2, ..., X_{s-1} = x_{s-1}) = P(X_s = x_s|X_{s-1} = x_s - 1) \tag{3.1}$$

More formally, **Eq. 3.1** denotes that the basis for drawing the new sample depends only on the steps before it and is independent of all previous steps. This embodies the limited memory concept that makes it a Markov Chain. We implement the Markov Chain Monte Carlo process using the Metropolis-Hastings algorithm after Gong et al. (2011) and Zhukovsky et al. (2016). In this algorithm, since the exact posterior distribution is unknown, we must draw from proposal distributions instead. The algorithm requires symmetrical proposal distributions in to exclude the effects of the chosen proposal distributions as demonstrated in **Eq. 3.2** (See also: **Table 3.1**). For each new sample in the chain, there is a chance the sample will be accepted and a chance the sample will be rejected. In order to determine the acceptance or rejection of a sample, we calculate the acceptance ratio, α and compare it to a random variate from the uniform distribution between 0 and 1.

$$\begin{aligned} \alpha &= \min\left(1, \frac{\pi(\theta_{proposal}|y)q(\theta_{t-1}|\theta_{proposal})}{\pi(\theta_{t-1}|y)q(\theta_{proposal}|\theta_{t-1})}\right) \\ &= \min\left(1, \frac{\pi(\theta_{proposal}|y)}{\pi(\theta_{t-1}|y)}\right) \\ &= \min\left(1, \frac{f(y|\theta_{proposal})\pi(\theta_{proposal})}{f(y|\theta_{t-1})\pi(\theta_{t-1})}\right) \end{aligned} \tag{3.2}$$

Table 3.1: Selected terms from **Eq. 3.2** and their descriptions

Term	**Description**
$\pi(\theta_{proposal}\|y)$	Posterior probability of observing a certain proposal given y
$q(\theta_{proposal}\|\theta_{t-1})$	Proposal probability of $\theta_{proposal}$ given θ_{t-1}
$\pi(\theta_{proposal})$	Probability of observing $\theta_{proposal}$
$f(y\|\theta_{proposal})$	Probability of observing production, y, given $\theta_{proposal}$
θ_{t-1}	Sample from previous step in Markov Chain
$\theta_{proposal}$	Proposed sample

When symmetrical proposal distributions are chosen, $q(\theta_{t-1}|\theta_{proposal}) = q(\theta_{proposal}|\theta_{t-1})$, leading to the simplification of the acceptance ratio in **Eq. 3.2** (See also: **Table 3.1**). One of the more desirable properties of Markov Chain Monte Carlo simulations is their ability to jump out of local minima or maxima, as applicable. This is a function of the acceptance ratio: when the ratio is calculated, if the new sample is more likely than the previous sample in the chain, the acceptance ratio is equal or greater to 1 and is automatically accepted. If the new sample is less likely than the previous sample, as measured by the objective function of choice, then the acceptance ratio is less than 1 but the sample will still be accepted if the random variate drawn is less than the acceptance ratio. In this way, the simulation has the ability to jump out of local minima or maxima of the posterior distribution while still tending to converge to a high likelihood region. The more likely the new sample is relative to the old one, the more likely it is to be accepted. An example of the Markov Chain Monte Carlo simulation converging to the high likelihood region of the 3D Rosenbrock function is shown in **Fig. 3.1** to illustrate the method.

We apply the Markov Chain process to Zhang et al.'s (2015) equation by selecting priors that allows the method to explore the space of reasonably plausible decline curves (Zhukovsky et al. 2016). Proposal distributions for each of the decline curve parameters are symmetric and either normal distributions or uniform distributions. Certain parameters, q_i,$\beta_e, \text{and}\beta_l$ are transformed onto a log scale for the prior and proposal distributions.

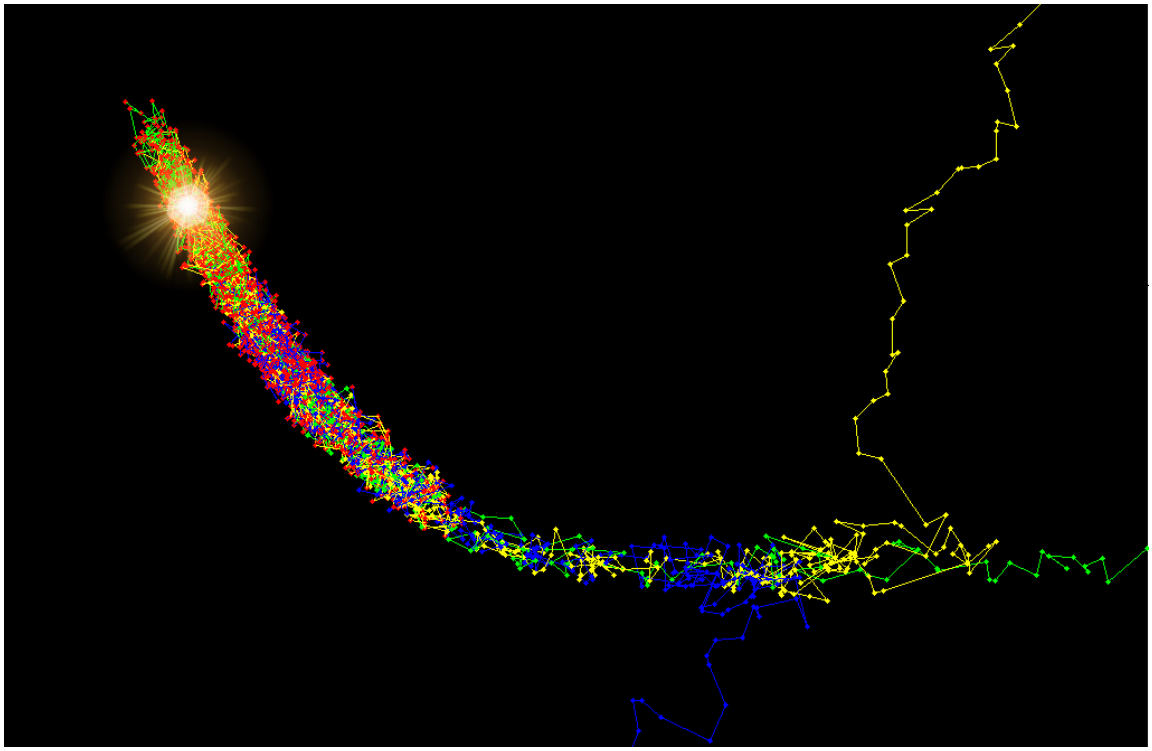

Figure 3.1: The Result of Three Markov Chains Running on the 3D Rosenbrock Function Using the Metropolis-Hastings Algorithm. The Algorithm Samples from Regions Where the Posterior Probability is High and the Chains Begin to Mix in These Regions (Wikipedia 2016)

The n parameter is on a linear scale. The general work flow follows **Fig. 1.1**.

In our work flow, each well is processed sequentially, starting with a least squares fit which serves as the beginning of the Markov Chain of decline curve parameters. The least squares fit was performed using MATLAB's nonlinear solver, using the constraints in **Table 3.1**. We returned the results of the least-squares fit to the main method in order to start the Markov Chain.

Once the least squares fit is performed, the method then proceeds to generate the Markov Chain, one sample at a time. Each new sample is accepted or rejected sequentially and the appropriate values are stored in the Markov Chain. The simulation continues to iterate

Table 3.2: Solver constraints for Zhang et al.'s (2015) decline curve parameters

Parameter	**Constraint**
β_e	0.01 <= β_e <= 1.5
β_l	0.01 <= β_l <= 0.15
n	0 <= n <= 2.5
q_i	0 <q_i <3,300,000 STB/mo

until it reaches the limit of 100,000 iterations. After the Markov Chain is constructed, we extract useful distributions of interest. Not only do we already have distributions for the decline curve parameters, q_i, β_e, β_l, and n, but we can also construct distributions of EUR and production at each month using the Markov Chain of decline curves. Once the Markov Chain is generated for a specific well, we use the output distributions of EUR and production at each forecasted month, which we use for area aggregation and areal hindcasting.

We also examine how the Markov Chain Monte Carlo process results in a narrower posterior distribution, as compared to the prior distribution, for each parameter (**Fig. 3.2**). Looking at examples of posterior distributions for the decline curve parameters for the Warren B well in the Eagle Ford (**Fig. 3.3**), we observe that the distributions have a lower variance than the prior distributions. By utilizing the method to create a Markov Chain of decline curve parameters, we incorporate new information from the production data and narrow the variance of each distribution from prior to posterior distribution.

Often, when introducing a new decline curve model or probabilistic method that involves production data, the issue of validation and calibration arises. If all of the available data is used in the input phase, there is none left to use to asses the accuracy of the predictions made by the method. Hindcasting is the practice of reserving data from the input phase and using it for validation purposes. In our method, we use only the first 12 months for input purposes, reserving the rest for hindcasting. In order to asses the accuracy of

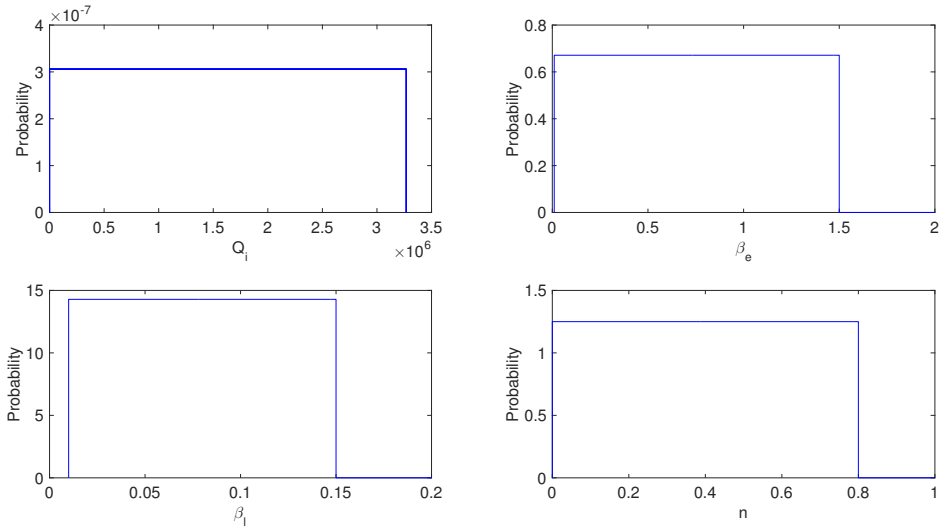

Figure 3.2: Prior Distributions for the Warren B Well in the Eagle Ford. The Prior Distributions are Deliberately Non-Informative So as to Allow the Markov Chain to Explore Plausible Combinations of Decline Curve Parameters.

our method, we must measure how often the predicted P90-P10 interval brackets the production reserved for the hindcasting on each well. An accurately calibrated probabilistic method can be expected to have a coverage rate of close to 80% for a P90-P10 interval of the "true reserves," the sum of the hindcasting production data.

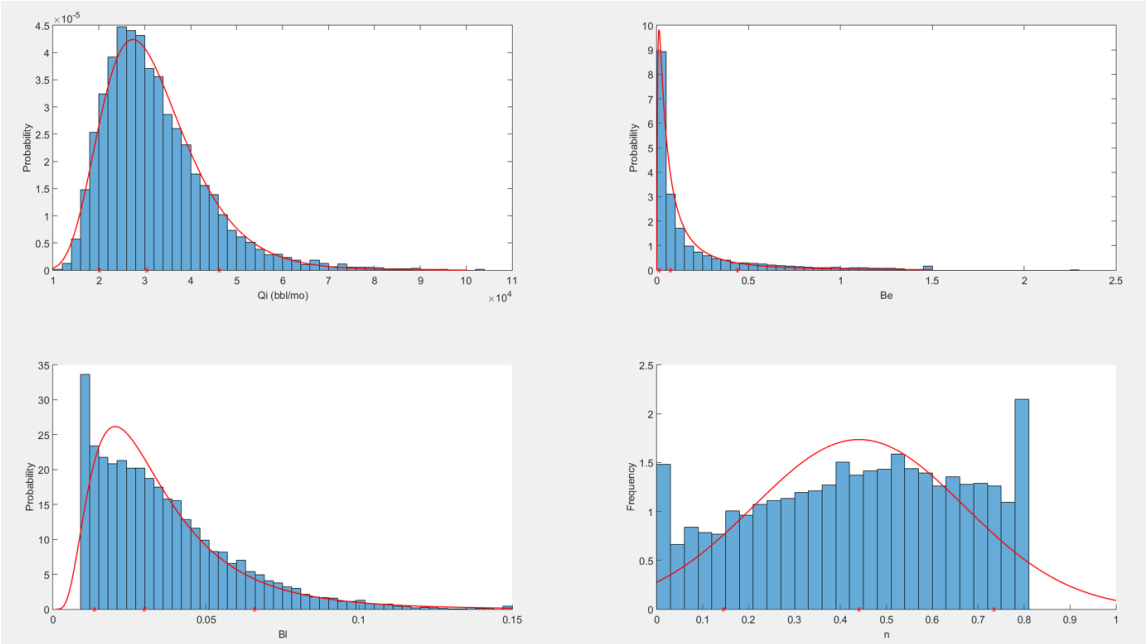

Figure 3.3: Posterior Distributions for the Warren B Well in the Eagle Ford. The Distributions for Most Parameters are Well Defined, While the Posterior Distribution for the n Parameter is Less So.

IV. APPLICATION OF METHOD TO EAGLE FORD FIELD DATA*

We now turn to the discussion of the application of our method to a large data set of wells in the Eagle Ford shale. We will discuss the specifics of applying the method, the results, the implications of data allocation, and limitations of the method.

4.1 Well Filtering Methodology

We apply our method to a selection of oil wells in the Eagle Ford oil window. We focus on one of the major areas of development in the oil window, namely, Karnes, Gonzaels, and DeWitt counties in the state of Texas (see **Fig. 4.1** for map of the Eagle Ford shale play). We start by obtaining from DrillingInfo the oil, gas, and water rates for all the wells in Karnes, Gonzales, and DeWitt counties that have been identified as drilled in the Eagle Ford with more than 3 years of production history as of August, 2015: more than 1,100 wells. We then proceeded to filter the wells as follows:

- Removed wells that did not meet minimum production requirements for bulk of well life
  - > 200 STB/mo. oil production
  - > 1,000 MSCF/mo gas production
- Removed wells with greater than 2 months of interruption to production data reporting
- Removed wells with less than 3 years of production history

*. Part of this chapter is adapted with permission from "Uncertainty Quantification in the EUR of Eagle Ford Shale Wells Using Probabilistic Decline-Curve Analysis with a Novel Model" by Isaac Zhukovsky, Ruben Mendoza, Michael King, and W. John Lee, 2016, Presented at Abu Dhabi International Petroleum Exhibition & Conference, 7-10 November, Abu Dhabi, UAE. Copyright 2016 by Society of Petroleum Engineers.

- Removed wells that experienced extensive intervention or drastic changes to production rates that appear to be man-made disruptions

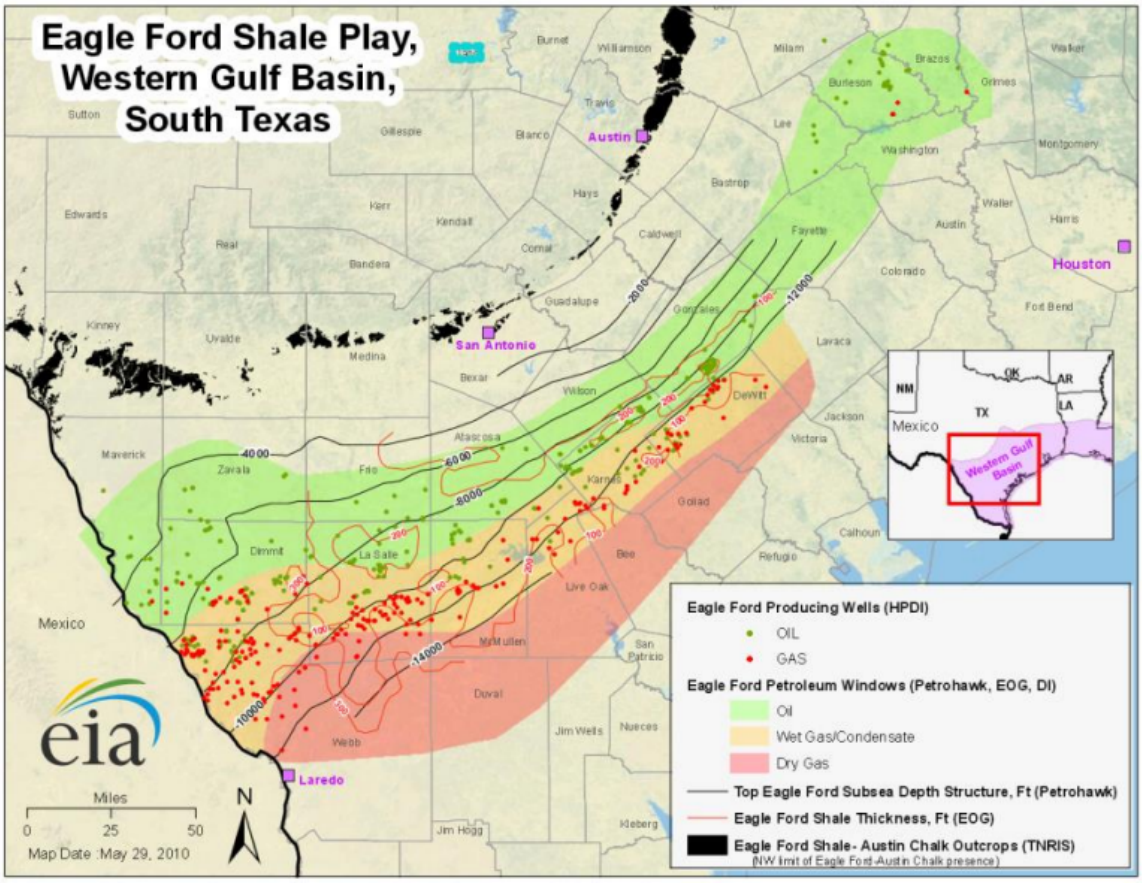

Figure 4.1: A Map of the Eagle Ford Shale Play that Delineates the Dry Gas, Wet Gas/Condensate, and Oil Windows. The Eagle Ford Stretches from Central Texas in Brazos County Southwest to the Border with Mexico. The Eagle Ford is Known in Mexico by a Different Name (EIA 2014)

In the end, we filtered the wells for length and tried to identify wells on "natural" decline. Wells that experienced re-stimulations, multiple cleanouts, or were subjected to extensive field operations were removed from the data set. We also tried to avoid wells

that had artifacts in their production histories as a result of production allocation necessitated by Texas's regulatory framework. While we attempted to minimize the impact of production allocation, it was impossible to avoid entirely. We will discuss the effect it had on the results of the analysis in **Section 4.4**. After the filtering process is complete, we proceed with 258 Eagle Ford oil wells.

4.2 Markov Chain Monte Carlo Processing

Once we have the filtered data set, we manipulate the format, but not the contents, of the dataset in Excel to prepare it for import into MATLAB. Each well is processed sequentially, and after the Markov Chain is constructed for each well, we extract P90, P50, and P10 EURs by fitting a distribution to the EUR data for the Markov Chain. We also obtain the distribution of forecasted production for the Markov Chain at each month and fit a distribution to it, so that we can extract the P90, P50, and P10 of the forecasted production at each month. This data is used to construct the P90, P50, and P10 decline curves. The calculation time for each well is between 5 and 10 seconds.

4.3 Hindcasting and Results

Once each well in the dataset has gone through the Markov Chain Monte Carlo process, we calculate the hindcasting coverage for the area. To do this, we sum the production data beyond 12 months for every well to form a hindcast production total, and check to see if it falls within the bounds of the P90 and P10 forecasts for the same interval. If a well falls within the P90-P10 interval, we considered it covered and add it to our correctly hindcasted well total. We divide by the total wells forecasted to obtain the hindcasting coverage rate for the area. For the 258 wells in our data set, the hindcasting coverage rate was 78.4% We obtain the P90, P50, and P10 EURs for all the wells in the data set and plot the data in histograms with fitted distributions. The data can be seen in **Figs. 4.2, 4.3, and 4.4**, for the P90, P50, and P10 EURs, respectively.

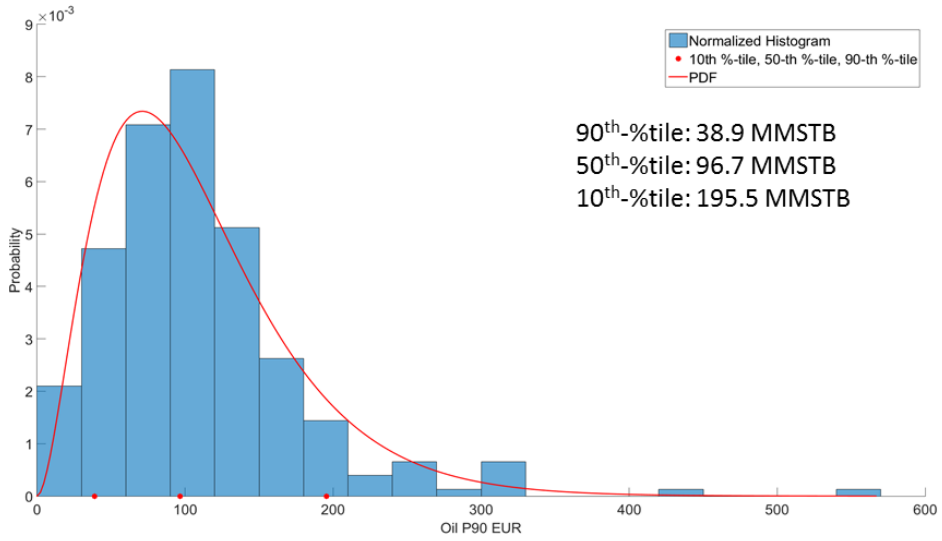

Figure 4.2: PDF of P90 EUR Distribution for the Eagle Ford Data Set. 90th %-tile 34 MSTB/well, 50th %-tile 121 MSTB/well, 10th %-tile 301 MSTB/well (Zhukovsky et al. 2016)

We plot the PDFs (without the histograms) together on same graph (**Fig. 4.5**), so that it is easier to see the way the shapes of the distributions change as we move through the probability percentiles.

We also plot the data as a CDF in **Figs. 4.6, 4.7, and 4.8**, for the P90, P50, and P10 EURs, respectively.

Finally, we plot the CDFs (without accompanying histograms) in **Fig. 4.9** in order to demonstrate how the CDFs change as we move through the various EUR distributions. A summary table of the various EUR distributions and their percentiles are complied in **Table 4.1**.

We can make the following observations:

- The distributions have a higher variance as we move from P90 $\rightarrow$ P50 $\rightarrow$ P10
- The EUR distributions are all log-normally distributed

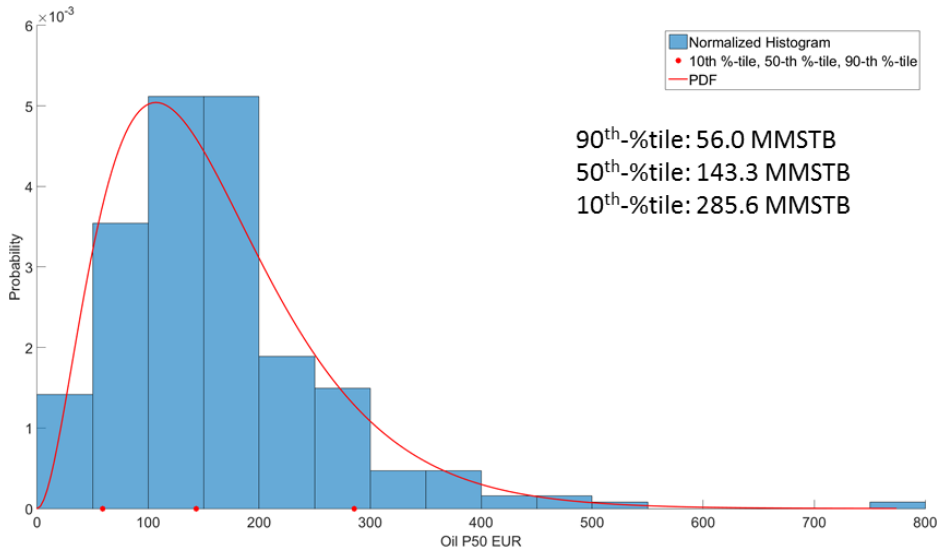

Figure 4.3: PDF of P50 EUR Distribution for the Eagle Ford Data Set. 90th %-tile 57 MSTB/well, 50th %-tile 193 MSTB/well, 10th %-tile 467 MSTB/well (Zhukovsky et al. 2016)

Table 4.1: Summary of the various percentiles of the P90, P50, and P10 EUR distributions

	90th-%tile EUR	**50th-%tile EUR**	**10th-%tile EUR**
P90 EUR Distribution	38.9 MMSTB	96.7 MMSTB	195.5 MMSTB
P50 EUR Distribution	56.0 MMSTB	143.3 MMSTB	285.6 MMSTB
P10 EUR Distribution	85.1 MMSTB	212.7 MMSTB	431.2 MMSTB

- The distributions become more skewed right as we move from P90 $\rightarrow$ P50 $\rightarrow$ P10

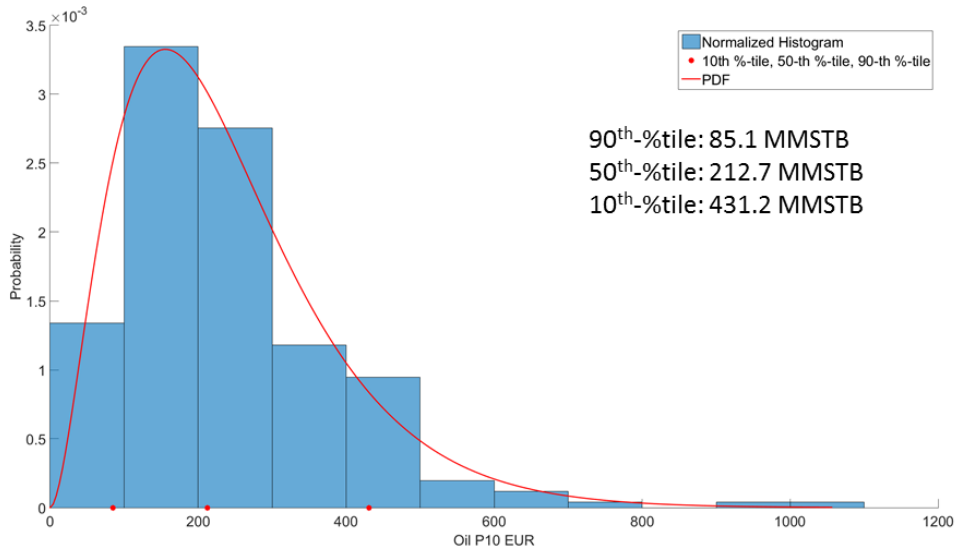

Figure 4.4: PDF of P10 EUR Distribution for the Eagle Ford Data Set. 90th %-tile 84 MSTB/well, 50th %-tile 279 MSTB/well, 10th %-tile 927 MSTB/well (Zhukovsky et al. 2016)

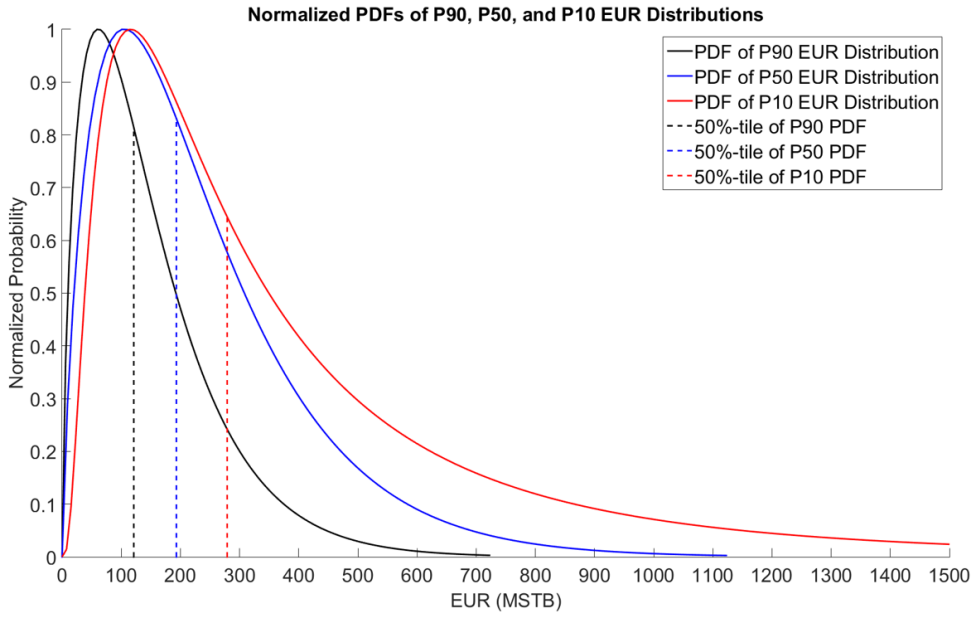

Figure 4.5: P90, P50, and P10 PDFs for the Eagle Ford Data Set. As We Move Through the Percentiles, the PDF of the EUR Distribution Becomes More Skew Right. All Three are Log-Normally Distributed (Zhukovsky et al. 2016)

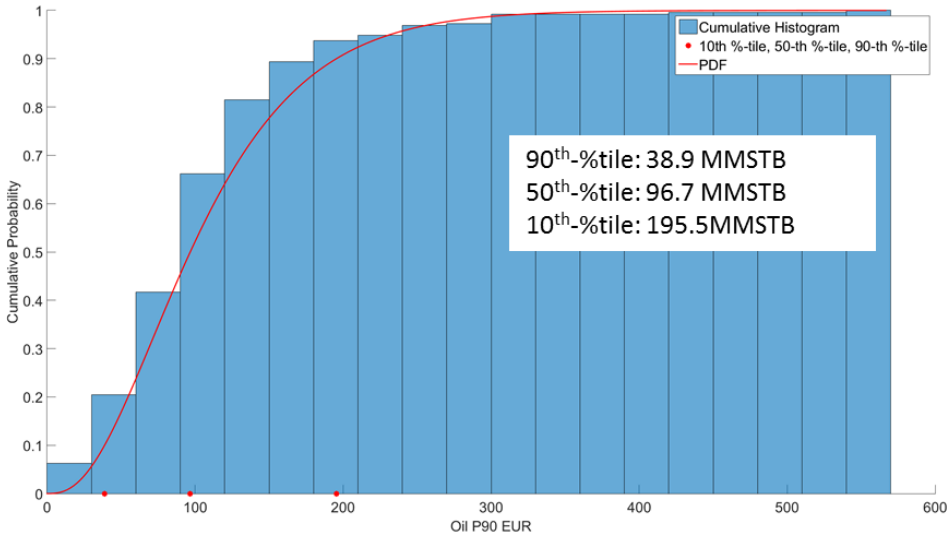

Figure 4.6: CDF of P90 EUR Distribution for the Eagle Ford Data Set. 90th %-tile 34 MSTB/well, 50th %-tile 121 MSTB/well, 10th %-tile 301 MSTB/well (Zhukovsky et al. 2016)

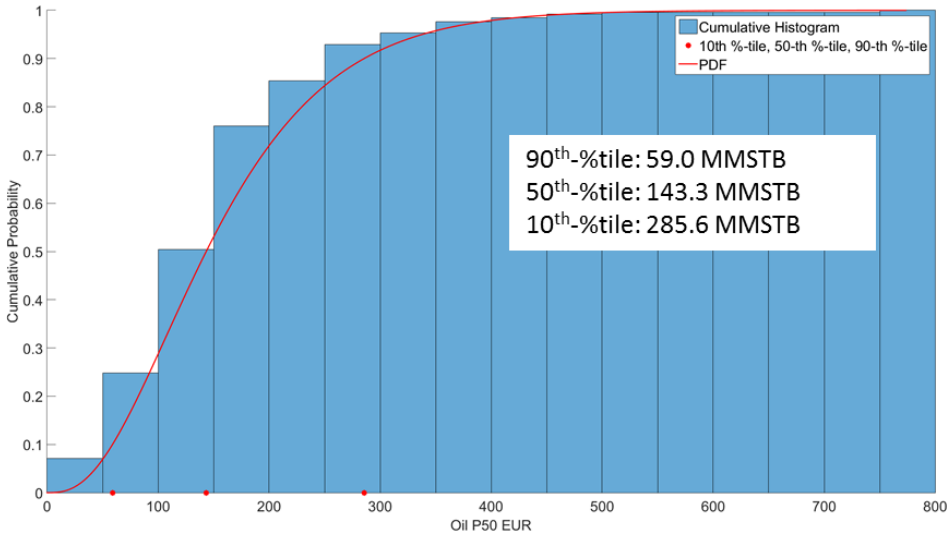

Figure 4.7: CDF of P50 EUR Distribution for the Eagle Ford Data Set. 90th %-tile 57 MSTB/well, 50th %-tile 193 MSTB/well, 10th %-tile 467 MSTB/well (Zhukovsky et al. 2016)

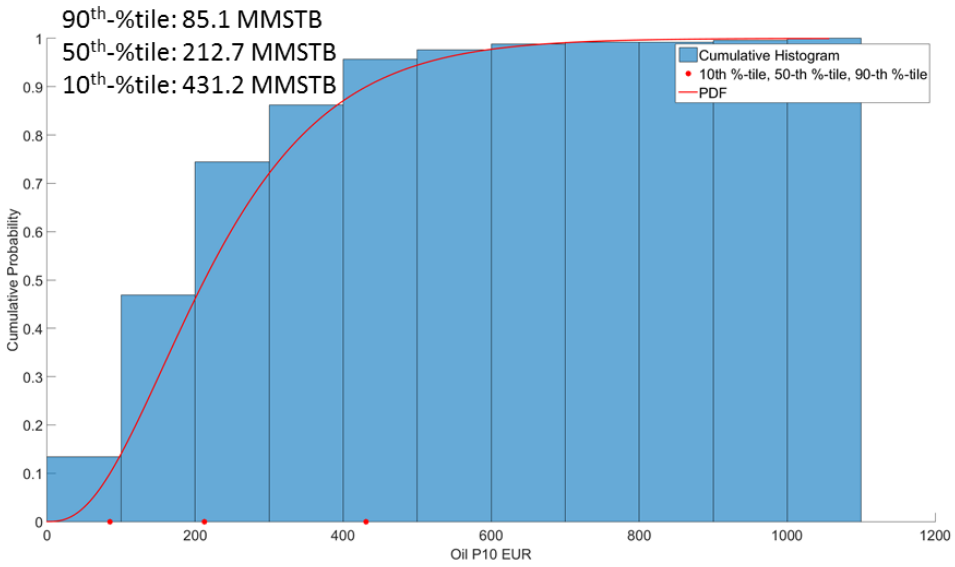

Figure 4.8: CDF of P10 EUR Distribution for the Eagle Ford Data Set. 90th %-tile 84 MSTB/well, 50th %-tile 279 MSTB/well, 10th %-tile 927 MSTB/well (Zhukovsky et al. 2016)

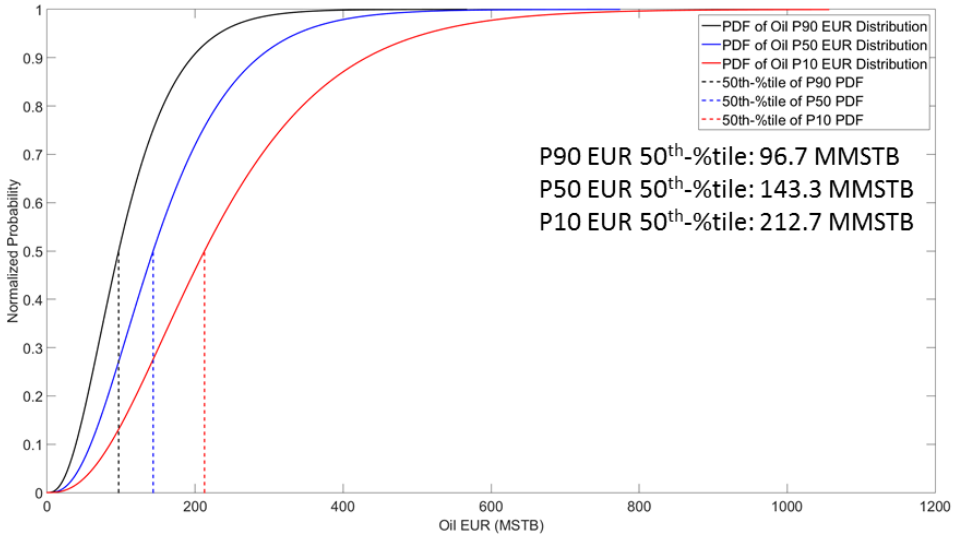

Figure 4.9: P90, P50, and P10 CDFs for the Eagle Ford Data Set. As We Move Through the Percentiles, the CDF of the EUR Distribution Becomes More Skew Right. All Three are Log-Normally Distributed (Zhukovsky et al. 2016)

4.4 Distribution of Error and Implications of Data Allocation

We now proceed to the error analysis phase. In order to test the validity of our results, we calculate the quantity (True EUR - P90)/(P10-P90), where the True EUR is the sum of the remaining production data available, for each well and plot a histogram of the numbers for the entire data set in **Fig. 4.10** (Zhukovsky et al. 2016). We observe that the distribution is centered around 0.5 and mostly symmetrical, indicating that the method is largely unbiased. However, the presence of a tail on the right hand side of **Fig. 4.10** indicates that the method is slightly biased towards the P90 EUR distribution. We also calculate several measures of error and compare them in **Table 4.4**. We observe that while the coverage rate of true reserves is quite good at 78.4% for an 80% confidence interval, the other measures report a significant amount of error.

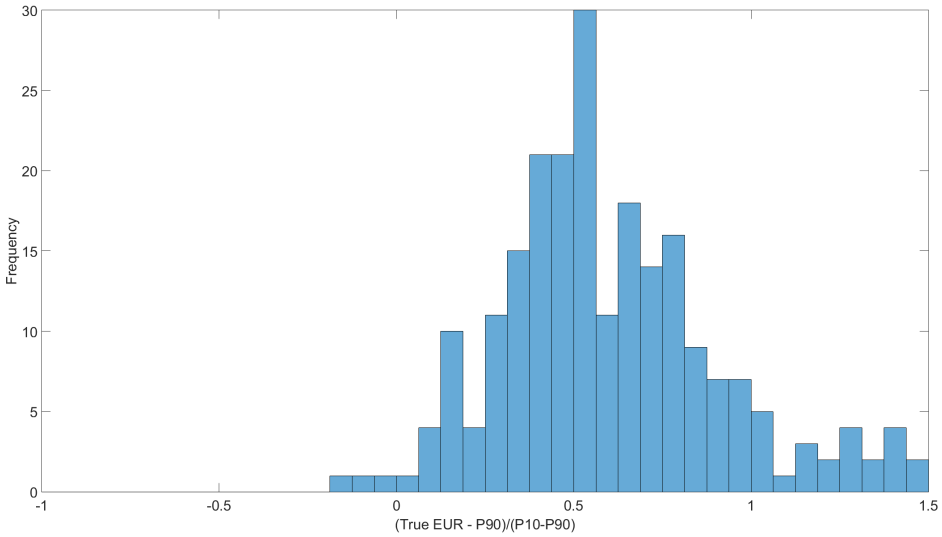

Figure 4.10: Distribution of Relative Error for the Markov Chain Monte Carlo Method. Error is Roughly Normally Distributed with a Center Around 0.5, with a Significant Tail on the Right Hand Side (Zhukovsky et al. 2016)

Coverage Rate of True Reserves for 80% Confidence Interval	78.4%
Error in True Reserves	27.5%
Average Relative Error (P50-True EUR)/(True EUR)	27.6%
Average absolute Relative Error Abs (P50-True EUR)/(True EUR)	37.8%

On an area basis, we can see that there is a significant deviation of the total from the true reserves. This implies that while we are predicting the bounds correctly based on early production data and accurately quantifying the uncertainty, there is still a significant deviation from the actual true reserves produced. We believe that the errors are caused by the way public data is allocated in the state of Texas (Zhukovsky et al. 2016). Each of the purveyors of publicly available production data obtain their production data directly from the Texas Railroad Commission, with production reported at a lease level. Since Texas does not use the Bureau of Land Management's Township-Section-Range system, there are large swaths of the state where leases are unitized to the square mile section standard and are highly variable in shape and/or size. This means that a given lease can contain a variable number of wells and be several square miles large with thousands of acres or contain just one well drilled in one formation. Public data aggregations must therefore allocate data from the lease level totals back to individual well production figures. The Texas Railroad Commission mandates that a well test is performed when a new well is brought online and that each producing well is tested once every twelve months. Using these publicly available well tests as a measure of relative well productivity, aggregators, like DrillingInfo, can allocate a portion of the lease production data to each individual well. The exact formula each company uses is a trade secret, so it is not always clear exactly which methodology is used. In addition to the fact that well tests are performed infrequently, the fact that the Railroad Commission does not mandate a standard well testing format complicates matters further. Finally, we observe that when wells are not brought

online at the same time during the year, allocation becomes even more difficult. All of these issues lead to unnatural step changes in production data when allocation factors change, an example of which is shown in **Fig. 4.11**.

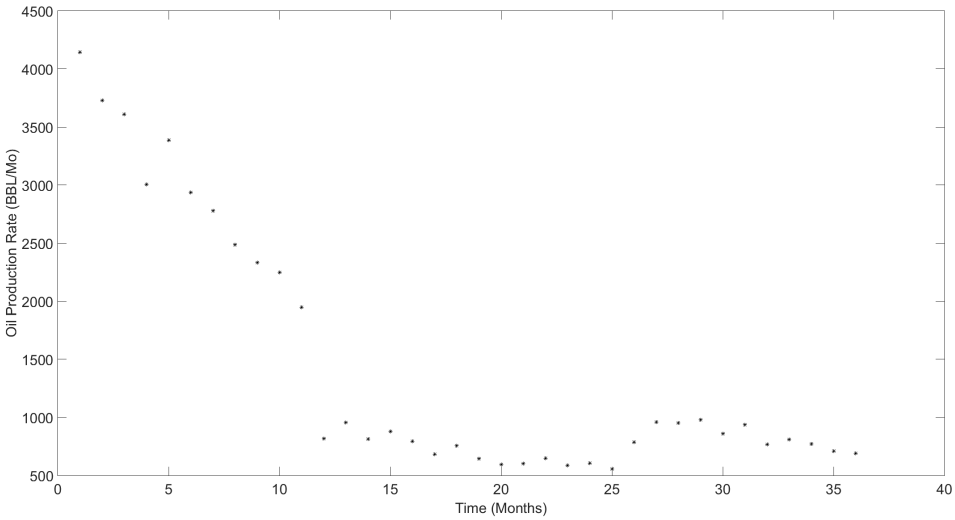

Figure 4.11: Effect of Changing Allocation Factor on Public Production Data. There are Several Shifts, with the Most Significant Occurring Around 10 Months (Zhukovsky et al. 2016)

The changes in allocation factor can lead to frequent shifts in production, representing a source of artificial noise in the data. We believe that this noise leads to difficulty in forecasting with only limited data available and adds to the error in the forecasts. The good news is that this issue can be eliminated by switching to a data source that is not subject to allocation or is subject to more accurate allocation, such as internal company data. There are also single well leases in Texas, but finding a significant concentration of single well leases in any given formation can be a matter of luck.

4.5 Limitations

While the coverage of the true reserves is excellent, with 199 out of the 258 wells correctly bracketed by the P90 and P10 predictions, the absolute and relative errors are significant at 38% and 28% respectively (Zhukovsky et al. 2016). The most likely source of this error is artificial allocation noise in the public data set used in this work. Unfortunately, the pubic data is what we have access to for this area of the Eagle Ford Shale.

While our Markov Chain Monte Carlo method works well for this decline curve model, in more complicated models it may be advisable to use a more sophisticated probabilistic method to quantify uncertainty. Certain methods are less strictly iterative and have more parallelization built in, such as the CURE clustering algorithm. This algorithm, presented by Guha, Rastogi, and Shim (1998), partitions probability distributions while simultaneously sampling the proposal distributions, increasing computation efficiency for complex systems.

Finally, while our method as coded is rather fast with a calculation time of 5-10 seconds per well, we acknowledge that reimplementing the method in a open source language like Python or R would likely increase the computational speed and would reduce cost of implementation of the method by removing its dependency on MATLAB.

V. SUMMARY, CONCLUSIONS, AND RECOMMENDATION FOR FUTURE WORK

5.1 Summary

In this work, we proposed a new method for forecasting shale wells using Zhang et al.'s (2015) growing drainage volume decline curve in a Bayesian framework (Gong et al. 2011; Zhukovsky et al. 2016). The existing literature offered many novel decline models and probabilistic methods coupled with more tradition decline curves, but few examples of probabilistic methods coupled with novel decline models. Our method quantified uncertainty while utilizing an empirical model designed for shale wells. We validated the method using the Warren B well in the Eagle Ford shale.

We have successfully applied our method to 254 wells drilled in the Eagle Ford shale in Karnes, Gonzales, and DeWitt counties. The results demonstrated that our method accurately quantified uncertainty on an areal basis. We asserted that the use of production data that was not allocated from the lease level would improve the accuracy of the method.

The proposed method can be applied to any shale formation through the use of tuning methods, making it very general in nature. Additionally, our method has a quick calculation time and is flexible in the nature and quantity of output distributions, enhancing its utility above beyond that of simply forecasting EURs.

5.2 Conclusions

We state the following conclusions based on this work:

1. The current state of decline curve analysis in shales requires the use of cumbersome composite models or models based on invalid physical assumptions for shale wells

2. The proposed method qunatifies uncertainty and allows for output distributions of

many statistics of interest

3. The proposed method provides P90, P50, and P10 probabilistic EUR estimates and eliminates user bias from the decline curve analysis process

4. The proposed method offers a fast calculation time and can be used to reforecast many wells automatically as new data becomes available

5. By automating the forecasting process for many wells and eliminating user basis, the method allows the implementing engineer to run the method frequently with a reduced work burden and use the output for planning purposes more frequently than is necessitated by bi-annual reserve reporting

5.3 Recommendation for Future Work

Based on our results, we recommend the following promising areas for future research efforts:

1. Investigation of the sensitivity of the method to length of input data. Our work focused on the use of 12 months of input data, but different lengths can be used. As with other probabilistic methods, we would expect that additional data would decrease the amount of uncertainty while less would increase the amount of uncertainty. Additional work is needed to verify this supposition.

2. Implementation of the method using individual well production data that has not been allocated, such as internal company data. Allocation factors and the way in which they are calculated introduced a significant amount of noise into the production data. Use of internal data that has less artificial noise would likely increase the accuracy of the method.

3. Modeling of the link between probabilistic EURs and potential causative factors such as completion parameters and geology factors, perhaps using a neural network to form the model. We believe that the use of neural networks is especially promising when coupled with our probabilistic EUR method. Many models have been proposed that purportedly link geology and completion inputs to deterministic EURs, but we believe that accurately quantifying uncertainty and incorporating it in the model definition may improve the accuracy and utility of these models.

REFERENCES

Anderson, D. M., Nobakht, M., Moghadam, S., et al. 2010. Analysis of Production Data from Fractured Shale Gas Wells. Presented at SPE Unconventional Gas Conference, 23-25 February, Pittsburgh, Pennsylvania, USA. SPE-131787-MS. http://dx.doi.org/10.2118/131787-ms.

Arps, J. 1945. Analysis of Decline Curves. *Transactions of the AIME* **160** (01): 228–247. http://dx.doi.org/10.2118/945228-g.

Capen, E. 1976. The Difficulty of Assessing Uncertainty (includes associated papers 6422 and 6423 and 6424 and 6425). *Journal of Petroleum Technology* **28** (08): 843–850. http://dx.doi.org/10.2118/5579-pa.

Cheng, Y., Wang, Y., McVay, D., et al. 2010. Practical Application of a Probabilistic Approach to Estimate Reserves Using Production Decline Data. *SPE Economics & Management* **2** (01): 19–31. http://dx.doi.org/10.2118/95974-pa.

Duong, A. N. 2011. Rate-Decline Analysis for Fracture-Dominated Shale Reservoirs. *SPE Reservoir Evaluation & Engineering* **14** (03): 377–387. SPE-137748-PA. http://dx.doi.org/10.2118/137748-pa.

EIA. 2014. *Updates to the EIA Eagle Ford Play Maps.* US Department of Energy: Energy Information Administration. https://www.eia.gov/maps/pdf/eagleford122914.pdf.

Fetkovich, M. 1980. Decline Curve Analysis Using Type Curves. *Journal of Petroleum Technology* **32** (06): 1065–1077. http://dx.doi.org/10.2118/4629-pa.

Freeborn, R. and Russell, B. 2012. How To Apply Stretched Exponential Equations to Reserve Evaluation. Presented at SPE Hydrocarbon Economics and Evaluation Symposium, 24-25 September, Calgary, Alberta, Canada. SPE-162631-MS. http://dx.doi.org/10.2118/162631-ms.

Fulford, D. S., Bowie, B., Berry, M. E., et al. 2015. Machine Learning as a Reliable Technology for Evaluating Time-Rate Performance of Unconventional Wells. Presented at SPE Annual Technical Conference and Exhibition, 28-30 September, Houston, Texas. SPE-174784-MS. http://dx.doi.org/10.2118/174784-ms.

Fulford, D. S., Bowie, B., Berry, M. E., et al. 2016. Machine Learning as a Reliable Technology for Evaluating Time/Rate Performance of Unconventional Wells. *SPE Economics & Management* **8** (01): 23–39. http://dx.doi.org/10.2118/174784-pa.

Fulford, D. and Blasingame, T. 2013. Evaluation of Time-Rate Performance of Shale Wells using the Transient Hyperbolic Relation. Presented at SPE Unconventional Resources Conference Canada, 5-7 November, Calgary, Alberta, Canada. SPE-167242-MS. http://dx.doi.org/10.2118/167242-ms.

Gong, X., Gonzalez, R. A., McVay, D. A., et al. 2011. Bayesian Probabilistic Decline Curve Analysis Quantifies Shale Gas Reserves Uncertainty. Presented at Canadian Unconventional Resources Conference, 15-17 November, Calgary, Canada. SPE-147588-MS. http://dx.doi.org/10.2118/147588-ms.

Guha, S., Rastogi, R., and Shim, K. 1998. CURE. Presented at Proceedings of the 1998 ACM SIGMOD international conference on Management of data - SIGMOD '98. Association for Computing Machinery (ACM).

Ilk, D., Rushing, J. A., Perego, A. D., et al. 2008. Exponential vs. Hyperbolic Decline in Tight Gas Sands: Understanding the Origin and Implications for Reserve Estimates Using Arps Decline Curves. Presented at SPE Annual Technical Conference and Exhibition, 21-24 September, Denver, Colorado, USA. SPE-116731-MS. http://dx.doi.org/10.2118/116731-ms.

Jochen, V. and Spivey, J. 1996. Probabilistic Reserves Estimation Using Decline Curve Analysis with the Bootstrap Method. Presented at SPE Annual Technical Conference and Exhibition, 6-9 October, Denver, Colorado, USA. SPE-36633-MS. Denver, CO. http://dx.doi.org/10.2118/36633-ms.

Johnson, N. L., Currie, S. M., Ilk, D., et al. 2009. A Simple Methodology for Direct Estimation of Gas-in-place and Reserves Using Rate-Time Data. Presented at SPE Rocky Mountain Petroleum Technology Conference, 14-16 April, Denver, Colorado, USA. SPE-123298-MS. Denver, CO. http://dx.doi.org/10.2118/123298-ms.

Lee, W. J. and Sidle, R. 2010. Gas-Reserves Estimation in Resource Plays. *SPE Economics & Management* **2** (02): 86–91. SPE-130102-PA. http://dx.doi.org/10.2118/130102-pa.

Panda, M. N. and Lake, L. W. 1994. Estimation of Single-Phase Permeability from Parameters of Particle-Size Distribution. *AAPG Bulletin* **78** (07). http://dx.doi.org/10.1306/a25fe423-171b-11d7-8645000102c1865d.

Rushing, J. A., Perego, A. D., Sullivan, R., et al. 2007. Estimating Reserves in Tight Gas Sands at HP/HT Reservoir Conditions: Use and Misuse of an Arps Decline Curve Methodology. Presented at SPE Annual Technical Conference and Exhibition, 11-14 November, Anaheim, California, USA. SPE-109625-MS. http://dx.doi.org/10.2118/109625-ms.

Valko, P. P. 2009. Assigning value to stimulation in the Barnett Shale: a simultaneous analysis of 7000 plus production hystories and well completion records. Presented at SPE Hydraulic Fracturing Technology Conference, 19-21 January, The Woodlands, Texas, USA. SPE-119369-MS. http://dx.doi.org/10.2118/119369-ms.

Valko, P. P. and Lee, W. J. 2010. A Better Way To Forecast Production From Unconventional Gas Wells. Presented at SPE Annual Technical Conference and Exhibition, 19-22 September, Florence, Italy. SPE-134231-MS. http://dx.doi.org/10.2118/134231-ms.

Wikipedia. 2016. Metropolis-Hastings algorithm. (December 6, 2016 revision), https://en.wikipedia.org/wiki/Metropolis%E2%80%93Hastings_algorithm. (accessed 14 February 2017).

Zhang, H., Cocco, M., Rietz, D., et al. 2015. An Empirical Extended Exponential Decline Curve for Shale Reservoirs. Presented at SPE Annual Technical Conference and Exhibition, 28-30 September, Houston, Texas, USA. SPE-175016-MS. http://dx.doi.org/10.2118/175016-ms.

Zhukovsky, I. D., Mendoza, R. C., King, M. J., et al. 2016. Uncertainty Quantification in the EUR of Eagle Ford Shale Wells Using Probabilistic Decline-Curve Analysis with a Novel Model. Presented at Abu Dhabi International Petroleum Exhibition & Conference, 7-10 November, Abu Dhabi, UAE. SPE-183138-MS. http://dx.doi.org/10.2118/183138-ms.